UCRL-JRNL-211515

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

# A New Paradigm for Identification of Classes of High Energy Gamma-Ray Sources

D. F. Torres, O. Reimer

April 19, 2005

Astrophysical Journal Letters

## Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

# A new paradigm for identification of classes of high energy gamma-ray sources

Diego F. Torres[1] and Olaf Reimer[2]

## ABSTRACT

A large fraction of the expected number of source detections of the forthcoming observatory Gamma-ray Large Area Space Telescope (GLAST) will be initially unidentified. We argue that traditional methodological approaches to identify individual detections and/or populations of gamma-ray sources present procedural limitations. These limitations will hamper our ability to classify the populations lying in the anticipated dataset with the required degree of confidence, in particular for those for which no member has yet been detected convincingly with the predecessor experiment EGRET. Here we suggest a new paradigm for achieving the classification of gamma-ray source populations that is based on implementing an a priori protocol to search for theoretically-motivated candidates. It is essential that such paradigm will be defined before the data is unblinded, in order to protect the discovery potential of the sample. Key to the new procedure is a quantitative assessment of the confidence level by which new populations can be claimed to have been discovered. When needed, small number statistics is applied for population studies in gamma-ray astronomy. Although we refer here explicitly only to the case of GLAST, the scheme we present can certainly be adapted to other experiments confronted with a similar combination of problems.

*Subject headings:* gamma rays: observations

## 1. Problem statement

The source wealth in the anticipated observations of the upcoming gamma-ray mission GLAST, potentially yielding the discovery of thousands of new high-energy sources following

---

[1]Lawrence Livermore National Laboratory, 7000 East Ave., L-413, Livermore, CA 94550, E-mail: dtorres@igpp.ucllnl.org

[2]Institut für Theoretische Physik IV, Ruhr-Universität Bochum, 44780, Germany. E-mail: olr@tp4.ruhr-uni-bochum.de

extrapolations from its predecessor experiment EGRET, will create a manifold problem for source identification. Catalogs of the most likely candidate sources (Active Galactic Nuclei -AGN-, neutron stars and pulsars) will not be complete to the required low flux levels in counterpart studies in the radio and X-ray waveband, predictably leaving many of these gamma-ray detections without plausible counterparts, and thus unidentified. But even if the pulsar and AGN catalogs were sufficiently deep or complete, they may not yield unambiguous source identifications. A legacy from the gamma-ray sources detected by EGRET is the indication that we are already failing to identify one or more new source populations, both at low and at high Galactic latitude. More specifically, the identification of variable, non-periodic, point-like sources at low galactic latitude, as well as of non-variable sources at high latitude is certainly still missing (Reimer 2001), all those exhibiting characteristics not obvious among the EGRET-detected pulsars or AGN. In addition, one has to bear in mind that a complete catalog for the suggested size of the populations would generate a total sky coverage of one or more candidates in every direction of the sky with a projected area similar to the instrumental point spread function (psf): There would be one or more AGN everywhere, and one or more pulsars in every low galactic latitude direction. This would make source identifications based solely on positional correlation meaningless, and impractical other identification scheme.

Whereas it is one of the goals of the GLAST mission to increase the number of blazar and pulsar identifications in order to enable scientific investigations of emission properties and evolution over a large parameter space, it is neither possible nor recommended -as we elaborate below- to strive for completeness of such catalogs of possible counterparts.

There has been a considerable number of definite gamma-ray source identifications achieved already, and certainly more will be known in coming years with the required high degree of confidence so as to claim that they pertain to the AGN or pulsar population. Thus, particularly interesting $\gamma$-ray source detections, initially lacking proper classification, can be pursued a posteriori. In contrast, the lack of completeness in pulsar and blazar catalogs will leave room allowing the discovery of other objects, if any, that may also be emitting high energy $\gamma$-rays at the current level of instrumental sensitivity. In the LAT era and beyond, if it is the objective to conclusively identify all gamma-ray source detections, we will predictably fail. This has to be seen in respect of the anticipated number of counterparts, their relatively faintness, the missing all-sky coverage in the relevant wavebands for counterpart studies, and expected ambiguities in densely populated regions of the gamma-ray sky. Therefore, we should aim to identify all classes of sources, not all individual detections, and subsequently attempt to gain in–depth astrophysical knowledge by studying the most interesting members among them.

Let us deepen into these statements.

The anticipated number of unidentified sources will preclude a individual deep multi-frequency study for every detected gamma-ray source, in the way it led to the identification of many blazars and, most exemplary, the Geminga pulsar.

Suppose that we have a sufficiently complete counterpart catalog, such that a member of it spatially coincides with most of the LAT sources. Does this imply that we have already identified all sources? To answer this question consider that we have, instead, a reasonably complete sky coverage of sources. Let us consider GRBs as an example. An overlay of all positional uncertainties for the GRBs reported in the most recent BATSE catalog covers the whole sky. Then, there is at least one GRB spatially coinciding with any possible counterpart. Consequently, a correlation analysis lacks identification capability, even when it is clear that not all populations of astrophysical objects are plausible candidates for GRB generation or hosting, nor that all of them should be probed. For instance, we can not claim, using correlation analysis, that GRBs have appeared more often in starbursts or luminous infrared galaxies than in normal galaxies.

Therein lies the dilemma. If the number of unidentified sources and/or the number of plausible candidates is sufficiently large, what will constitute an identification? How shall we found evidence for new populations of sources, and new members within these populations, in the GLAST era?

The identification scheme for $\gamma$-ray sources in the pre-GLAST era is primarily based upon multifrequency follow-ups, both in bottom-up and bottom-down approaches regarding photon energy. If there is a prediction of a periodic signal of the flux, that alone unambiguously label the source. This, however, will happen for only a fraction of LAT detections, either because of the absence of contemporaneous pulsar timing solutions, or because of the shortage of precise variability predictions for accretion powered X-ray binaries. Note that variability of $\gamma$-rays probes, generally, timescales, not periodicities, and can be used predominantly to *rule out* membership into classes and only when it has been found at a significant level. For example, if a given source *is* variable, we consequently assume that it is not produced in phenomena on timescales larger than the corresponding exposure. In essence, this will rule out all possible counterparts producing steady gamma-ray fluxes. However, if a theoretically compatible variability timescale exists, that in itself will not constitute more than a suggestive indication. Indeed, it will rather prompt the need of carrying out follow-up observations, which will necessarily require a considerable amount of time and resources, without guaranteed success of achieving an unique identification. The bottom line is that adopting this scheme, with LAT observations only, particularly during the first year of data taking, we may limit our capability to identify new populations of sources.

If we have a positional correlation between a candidate and an unidentified gamma-ray source, and in addition there exist a matching variability timescale between theoretical predictions for such object and the measured data, then how can we, with nothing else, definitely say that the identification was achieved? And even if we convince ourselves to assert it, is it sufficient to find one, two, or how many of such individual cases in order to claim the discovery of a new population of sources with satisfactory statistical significance? And, finally, how would the latter be quantitatively evaluated?

Not having *a priori of the expected number of source detections* a criterion by which to answer the previously posed problems will confront us with a situation of ambiguity between results achieved by applying different classification standards, with no instance to decide in a unbiased way whether an identification has been achieved or not.

Worst yet for seeking source identifications in the LAT era is the fact that most of the theoretically anticipated potential LAT sources are expected to present steady fluxes and spectra, such as the case of supernova remnants (e.g., Torres et al. 2003) or luminous infrared galaxies (e.g., Torres et al. 2004), or galaxy clusters (e.g., Reimer et al. 2003), or high latitude translucent molecular clouds (Torres et al. 2005), or others. We are returning to the domain of spatial associations in conjunction with a large number of unidentified gamma-ray sources and an increasingly imaginative number of possible counterparts. In order to overcome this predictable problem, a paradigm shift in the way we look into source classification is suggested. We need to define a sensitive and quantitative criterion, by which using a combination of the predictive power of theory and the spatial coincidences of counterpart candidates, we could identify both variable and non-variable populations. This Letter provides what we consider is a feasible scheme for defining such a criterion. Although we refer here explicitly only to the case of GLAST-LAT, the scheme we present, can certainly be adapted to other experiments confronted with a similar combination of problems, for example in neutrino astronomy.

## 2. Identification of $\gamma$-ray populations

Here we elaborate a scheme to identify and classify new $\gamma$-ray source populations.

### 2.1. What to search for?

Starting from a theoretical prediction of a population of sources to be detectable above the LAT instrumental sensitivity, for which individuals' coordinates are known, we propose to assume a

- *Theoretical censorship:* we request as part of the criterion that predictions (ideally of multiwavelength character) are available for a subset of the proposed class of counterparts.

This request is made to avoid the blind testing of populations that may or may not produce $\gamma$-rays, but for which no other than a spatial correlation result can be achieved a posteriori. If there is no convincing theoretical support that a population can emit $\gamma$-rays before executing the search, such population may not be sought this way. Although obvious, it should be explicitly stated that we will not, by applying this method, disallow the possibility of making serendipity discoveries. Imposing of a theoretical censorship is not just a matter of theoretical purity, but rather it is statistically motivated, as we explain below.

### 2.2. Protection of discovery potential

By probing a large number of counterparts candidates with at least equally large number of trials with the same data set, one *will* find positive correlations, at least as a result of statistical fluctuations. Then, to claim significance, one would have to check if the penalties that must be paid for such a finding (i.e., the fact that there were a number of trials that led to null results) does not overcome the significance achieved. Needless to say, a number of possible bias are expected to influence the computation of the penalties. The example here is ultra high energy cosmic rays (UHECRs), where there are already a number of dubious, correlation-based, discovery claims, even when the sample of events is small (see, e.g. Torres et al. 2003 for a discussion). GLAST-LAT, and in general $\gamma$-ray astronomy, can prepare to overcome this difficulty before embarking in the new era of photon wealth, as ultra high energy cosmic ray physics did before unblinding data from the Pierre Auger observatory (Pierre Auger Collaboration 2003). In this sense, this part of our criterion is rather similarly defined. We require an

- *A priori protocol:* The populations that are to be tested in the GLAST-LAT data must be defined before the initial data release.

Here, a *protocol* is basically a budget for testing correlations. Every test will consume part of this budget up to a point that, if we still proceed in testing, there can be no statistical significant detection claim achieved anymore. A protocol secures that a detection of a population can be made with confidence in its statistical significance for a number of interesting classes. As remarked by the Pierre Auger Collaboration (2003), when confronted with claims made in the absence of an a priori protocol, one may assume that a very large number of failed trials were made in order to find the positive results being reported, and thus disregard the claims altogether just by denying statistical weight. Otherwise stated, we might be asked for proof that the penalty for failed trials has been accounted for and is indeed below a required statistical significance. This may turn out to be, either very difficult to achieve or strictly impossible because of the possible biases in penalties definitions.

Additional thorough exploration of the same data set for expected or unexpected populations can (and will) certainly be made, although if the budget is spent, without the strength of immediate discovery potential. A positive additional search must be thought of as a way of pointing towards new populations of sources to be tested with additional or independent sets of data then, e.g., the second or fifth year LAT catalog. An specific, but evolving protocol should be determined a priori for each of these data sets in such a way we test all the populations we are interested in.

Summarizing, if using the same set of data, claiming the discovery of one population affects the level of confidence by which one can claim the discovery of a second. Then, suppose for definiteness that the *total* budget is a chance probability equal to $\mathcal{B}$, e.g., $10^{-4}$. That is, that a claim for population(s) discovery has to be better than one having a probability of chance occurrence equal to $\mathcal{B}$, and that we want to test $A$, $B$, $C$ ... classes of different sources (say, radio galaxies, starbirth galaxies, microquasars, pulsars, AGN, etc.). The total budget can then be divided into individuals, a priori, chance probabilities, $P_A$, $P_B$, etc., such that $\sum_i P_i = \mathcal{B}$. This implies that population $i$ will be claimed as detected in this framework if the a posteriori, factual, probability for its random correlation, $P^{\rm LAT}(i)$, is less than the a priori assigned $P_i$ (as opposed to be less only than the larger, total budget).

We could go a step forward and manage the budget of probabilities. For some populations, e.g., those which were not detected in EGRET observations, we can less confidently agree that they will be detected, or perhaps for some the number of their members may be low enough such that a detection of only several of its individuals would be needed to claim a large significance. In this situation, we would thus be less willing to spent a large fraction of the discovery budget in them, i.e., we would choose a relatively higher $P_i$, so that it would be easier to find $P^{\rm LAT}(i) < P_i$. For others, say AGN and pulsars, we know that they will be detected, but we can statistically prove it in this way, by assigning a very low $P_i$, in such a way to make harder for the test to pass. If one or more of the tests, i.e., is for several $i$-classes, $P^{\rm LAT}(i) < P_i$, is fulfilled, the results are individually significant. First, because we protected our search by the a priori establishment of the protocol (a blind test) and second, because the overall chance probability is still less than the total budget $\mathcal{B}$.

We refrain here to explicitly propose which are the populations to be tested and how large the a priori probability assigned to each of them as well as the exact number for the total budget $\mathcal{B}$ should be. This ultimately has to be carefully studied by the LAT collaboration, although obvious choices can be compiled and argued. We only emphasize that this should be done before the data is taken. Now we proceed towards a most delicate issue, that of quality control.

### 2.3. How to search and assess the quality of the findings

The last constituent of a methodological procedure to identify classes of new $\gamma$-ray sources is the application of

- *common quality assessments:* we urge that a strict statistical evaluation is mandatory before a claim of discovery of a new source population can be made. An objective method is presented in the following.

We start by assessing the number of members of the relevant candidate class being probed, for which predictions exist, that coincide with LAT source detections of unidentified gamma-ray sources. Let $\mathcal{C}(A)$ represent this number for population $A$. We recall that for EGRET, measures of the quality of the positional association were introduced motivated by the size of the EGRET likelihood contours (Mattox et al. 2001, Sowards-Emmerds et al. 2003). A similar assessment can be done for LAT. In what follows, for the sake of simplicity, we will assume that we deal with equally probable coincidences, when a projected position is less distant than, say, the 95% confidence contour.

Let $\mathcal{N}(A)$ be the number of known sources in the particular candidate population $A$ under analysis and $\mathcal{U}$ the number of LAT detections. Let $\mathcal{P}$ be the probability that in a random direction of the sky we find a LAT source. The probability $\mathcal{P}$ should take into account instrumental detectability issues (exposure gradients, imprecision of the diffuse emission model, etc.) as well as, at low Galactic latitudes, expected Galactic structures.

As an example which omits the latter complications, one may use angular coverage (the ratio between the area covered by $\mathcal{U}$ sources and that of the sky region upon which these sources are projected). In what follows, we will assume that such method is in place for LAT and that $\mathcal{P}$ can be computed for a given region of the sky. Note that to compute $\mathcal{P}$ we do not need any information about the candidates, but just some sensible extrapolation of the expected number of detections of sources that have been already identified. The value of $\mathcal{P}$ is obtained a priori of checking for any population.

Whatever the method, $\mathcal{P}$ is expected to be small for LAT. To give an example, if we take just a coverage assessment at high Galactic latitudes ($|b| > 10$), and we assume that there will be a thousand detections, and that the typical size of the error box of LAT sources is a circle of radius 12 arcmin, then $\mathcal{P} \sim 3 \times 10^{-3}$. At lower latitudes, we expect $\mathcal{P}$ to be between 1 to 2 orders or magnitude larger. We believe that a more careful treatment of source number predictions and the range of expected source location uncertainties will reduce the value of $\mathcal{P}$ from such simple estimations. Such low values for $\mathcal{P}$ make the product $\mathcal{P} \times \mathcal{N}(A)$ typical less than 1-10, for all different candidate populations. We will refer to this product as the *noise expectation*, i.e., this is the number of coincidences which one would expect even when there is no physical connection between the LAT detections and population $A$.

The number of excess detections above noise will be, $\mathcal{E}(A) = \mathcal{C}(A) - \mathcal{P} \times \mathcal{N}(A)$.[3] Two cases can be distinguished. The two largest populations of plausible candidates (pulsars and blazars) will also present the largest number of coincidences, since it is already proven that they do emit high energy $\gamma$-rays above LAT sensitivity. Let's assume that there are 2000 catalogued AGNs; with the quoted value of $\mathcal{P}$, all coincidences in excess than 6 are beyond the random expectation. The reality of the population in the EGRET catalog make us expect that $\mathcal{C}(\mathrm{AGN}) \gg 6$, and thus that the number of excesses would be equally large. In this case, we are in the domain of large number statistics and a probability for the number of excesses to occur by chance, $P^{\mathrm{LAT}}(\mathrm{AGN})$ can be readily computed.

A different case appears when the second term in the expression for $\mathcal{E}(A)$ is a small quantity. Two scenarios may be found: if the number of coincidences for that population is large compared with the noise, we are again in the domain of large number statistics, as in the case of AGN or pulsars. This will –most likely– not happen for many (or perhaps for any) of the new populations we would like to test. Thus, in general we are in the realm of small number statistics: we should test the null hypothesis for a new source population against a reduced random noise (see Feldman & Cousins 1998, also Gehrels 1986).

Let us analyze now an explicit example. We are testing a null hypothesis (e.g., X-ray binaries are not LAT sources). That is represented by 0 predicted signal events, i.e. total number of events equal to the background in Table 2-9 (see leftmost columns) of Feldman & Cousins (1998). Suppose for definiteness that $\mathcal{P} \sim 3 \times 10^{-3}$ and $\mathcal{N}(A)$ is equal to, say, 200, then the number of chance coincidences (the noise or background) is 0.5. Thus, if we find more than 5 individual members of this class (e.g. superseding the confidence interval

---

[3]Trivially, the previous formula for $\mathcal{E}$ shows, in correspondence with what we have discussed earlier, that if the number of sources is so large that $\mathcal{P} \to 1$, $\mathcal{E} = 0$. If instead, the number of members in the potential counterpart class is so large that $\mathcal{C}(A) \to \mathcal{P}\mathcal{N}(A)$, then $\mathcal{E} = 0$ too. In both cases, there is no way to distinguish whether the population is physically associated.

0.00-4.64) correlated with LAT sources, we have proven that the null hypothesis is ruled out at the the 95% CL; i.e., we have untangled the existence of a new $\gamma$-ray source population. [4]

Note that at this stage there is no variability analysis involved. If we were to add the search on compatible variability timescales, the confidence level of the detection will even improve. If, instead, we find no more than 2 individual sources in the same example, then we have no evidence by which to claim the existence of this population at that level of confidence. Using the small number statistics formalism, we can convert the level of confidence achieved for each population into a probability for random appearance, i.e., $P^{\rm LAT}({\rm A})$. Subsequently, a minimum requirement can be defined for the claim of detection of a new source population accordingly. This would ultimately allow to compare each of the latter values with the a priori budgeted $P_A$ and decide about detectability. Managing $P_A$ is equivalent to requesting different populations to appear with different, intelligently selected, levels of confidence.

If theorists' imagination were to bear resemblance with reality, note that by using this method, detecting just a few members of each class may allow to achieve significant levels of confidence, justified by the existence of the imposed theoretical censorship and protected by an a priori protocol.

## 3. Concluding remarks

The proposed criterion for identification of $\gamma$-ray source populations integrates three different parts:

- A *theoretical censorship* that secures that we know what we are looking for instead of just executing repeated searches that would reduce the statistical significance of any possible positive class correlation.
- An *a priori protocol* that protects the significance by which to claim the discovery of a number of important population candidates and gives guidelines as to how to manage the probability budget

[4]A confidence interval $[\mu_1, \mu_2]$ is a member of a set, such that the set has the property that $P(\mu \in [\mu_1, \mu_2]) = \alpha$, i.e., the probability to find $\mu$ in the interval $[\mu_1, \mu_2]$ is $\alpha$. Here $\mu_1$ and $\mu_2$ are functions of an observable $x$, and the previous equation refers to the varying confidence intervals $[\mu_1, \mu_2]$ from an ensemble of experiments with fixed $\mu$. For a Poisson process, when the observable $x$ is the total number of observed events $n$, consisting of signal events with expected mean $\mu$ (in the case of a null hypothesis, $mu = 0$), and background events with mean $b$, the probability of measuring $n$ given $\mu$ is $P(n|\mu) = (\mu + b)^n exp(-(\mu + b))/n!$ This can be used to translate a given number of coincidences into a probability for it to happen by chance.

- A *quality assessment* that assigns probabilities both in the large and in the small numbers statistical regime. For the latter, it is achieved by introducing that the hypothesis to test in searching for new population of sources is the null hypothesis against a reduced noise level.

It is useful to note that LAT will be in a privileged position to actually identify new population of sources. If LAT would have an additional order of magnitude better sensitivity, with no very significant improvement in angular resolution, a situation similar to the GRB case would appear, i.e., a flat distribution of unidentified sources with a few privileged individuals only which are extensively traced in multifrequency studies. Essentially, we would find a gamma-ray source coinciding with the position of every member of any population under consideration. And thus, we would lack the capability to achieve discoveries by correlation analysis. This is, perhaps, already indicating that a next generation high energy $\gamma$-ray mission after GLAST-LAT should not be exclusively sensitivity-driven if no significant improvement in angular resolution can be achieved.

We finally note that the potential of this methodological procedure is not limited to the anticipated cases explicitly discussed here. By applying the proposed scheme, one can also check spurious classifications in an objective way, and test subsamples among the expected classes of sources (e.g., FSRQs in correspondence of their peak radio flux, or BL Lacs in correspondence of their peak synchrotron energy, i.e. LBLs vs. HBLs).

The work of DFT was performed under the auspices of the U.S. D.O.E. (NNSA), by the University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48. OR acknowledges support by DLR QV0002. We thank Luis Anchordoqui for discussions.

## REFERENCES

Feldman G. J. & Cousins R. D. 1998, Physical Review D 57, 3873

Gehrels N. 1986, ApJ 303, 336

Mattox J. R., Hartman R. C. & Reimer O. 2001, ApJS 135, 155

Pierre Auger Collaboration (Roger W. Clay for the collaboration) 2003, Proceedings of 28th International Cosmic Ray Conference 421-424 e-Print Archive: astro-ph/0308494

Reimer O., Unidentified Gamma-Ray Sources in The Universe in Gamma Rays, Ed. V. Schönfelder Springer Astronomy & Astrophysics Library 2001, p.319-338 ISBN 3-540-67874-3

Reimer O., Pohl M., Sreekumar P. & Mattox, J. R. 2003, ApJ 588, 155

Sowards-Emmerd D., Romani R.W., & Michelson, P.F. 2003, ApJ 590, 109

Torres D. F., Reucroft S., Reimer O. & Anchordoqui L. A. 2003, ApJ 595, L13

Torres D. F., Romero G. E., Dame T. M., Combi J. A. & Butt Y. M. 2003b, Physics Reports 382, 303

Torres D. F., Reimer O., Domingo-Santamaría E. & Digel S. 2004, ApJ 607, L99

Torres D. F., Dame T. M. & Digel S. W. 2005, Astrophys. J. 621, L29

This preprint was prepared with the AAS LaTeX macros v5.0.